



**N1 FY10 Research Project**  
**Research Sponsor: Navy Recruiting Command**  
**POC: Mr. John Noble, Head of Research**



**Matching Navy Recruiting Needs with Social Network Profiles Using  
Lexical Link Analysis**

**Primary Researcher**

**Dr. Ying Zhao**

**408.252.7746**

**408.218.8484**

***Associate Researchers:***

***Dr. Douglas J. MacKinnon***

***Dr. Shelley P. Gallup***

***Research Associate Professors***

***Distributed Information Systems Experimentation (DISE)***

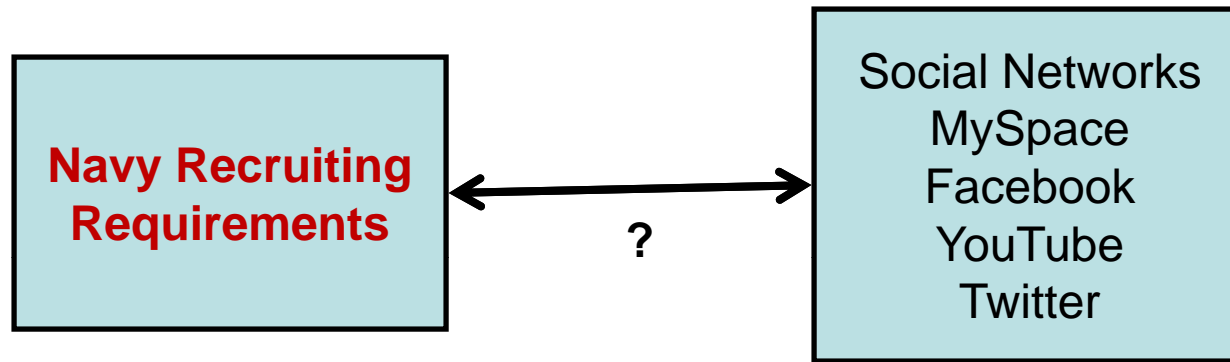
***Naval Postgraduate School***

Report Documentation Page				Form Approved OMB No. 0704-0188	
Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.					
1. REPORT DATE <b>JAN 2010</b>		2. REPORT TYPE		3. DATES COVERED <b>00-00-2010 to 00-00-2010</b>	
4. TITLE AND SUBTITLE <b>Matching Navy Recruiting Needs with Social Network Profiles Using Lexical Link Analysis</b>				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S)				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) <b>Naval Postgraduate School,Distributed Information Systems Experimentation (DISE),Monterey,CA,93943</b>				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION/AVAILABILITY STATEMENT <b>Approved for public release; distribution unlimited</b>					
13. SUPPLEMENTARY NOTES <b>Personnel and National Security: A Quantitative Approach (Unclass), 25-28 January 2010, Johns Hopkins University Applied Physics Laboratory, Laurel, Maryland</b>					
14. ABSTRACT					
15. SUBJECT TERMS					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT <b>Same as Report (SAR)</b>	18. NUMBER OF PAGES <b>25</b>	19a. NAME OF RESPONSIBLE PERSON
a. REPORT <b>unclassified</b>	b. ABSTRACT <b>unclassified</b>	c. THIS PAGE <b>unclassified</b>			



# Research Issues/Scope

## Matching Navy Requirements in a Social Networking Environment



## Research Questions

- Do current service entry profiles match well with candidate profiles in the social networking environments?
- What is the methodologies facilitate matching profiles that could potential inform Navy Recruiting Marketing decisions?
- Are there sufficient matches to make marketing in this environment cost effective?



# Proposed Method

- **Lexical Link Analysis (LLA)**
  - Features – terms/keywords to describe Navy recruiting needs and candidate profiles
  - Link the features in context of dynamic social network environments, learn from on-going market places
- **Previous applications of (LLA)**
  - System self-awareness/ program-awareness: DISE demonstrated methodology within the Navy acquisition community
    - Matched Navy needs (e.g. need statements) with technologies (e.g. Trident Warrior technologies)
    - Resulting in improved acquisition decisions
    - (See next slide for example)





# Navy Recruiting Research Tasks



## Select and Pre-Process data sets

- Obtain Facebook data samples - 1,700 Facebook profiles site
- Obtain service entry requirement descriptions and extract features for service entry requirements
- Pre-process and extract features from a sample data set

## Apply Proposed Approaches

- Profile the current service entry requirements using Lexical Analysis
- Profile Facebook sample data using Lexical Analysis
- Link profiles with rating descriptions using LLA generating potential recruiting candidates
- Develop mapping interface output matching qualified candidates

## Prepare Deliverables

- NPS Technical Report submitted at the completion of the project
- Presentation at Annual Navy Workforce Research & Analysis Conference in DC, May 2010
- A Trade Journal or Conference Paper submission in the area of Knowledge Management



# Relevant Publication

## Facebook Data Set



- “Tastes, Ties, and Time (T3): A new social network dataset using Facebook.com”
  - Authors
    - Kevin Lewisa, Jason Kaufmana, Marco Gonzaleza, Andreas Wimmerb, Nicholas Christakis
  - Affiliations
    - *Department of Sociology, Harvard University, United States*
    - *Department of Sociology, University of California, Los Angeles, United States*
- Note: T3 dataset ensures privacy of students’ profiles



# Published Research Data

- **1,700 Facebook profiles**
  - **Data set includes**
    - User profiles
    - Interests
    - Communities
    - Forums, group affiliations (schools, colleges and universities, companies, etc.)
    - Friends list, fandom (fan of)
    - Endorsements (supporter of)
- **Navy Enlisted Rating descriptions**
  - *Hard Cards*
  - Officer Community Descriptions





# Potential Research Benefits



- Improved strategic decisions
- Lexical links can be used to improve
  - Recruiting campaigns
  - Advertisements
  - Regionalization understanding
  - Marketing event strategies
- Real-time marketing
  - Facebook page that takes the feed and disseminates recruiting marketing messages



# Potential Impact



# Direct Advertising Using LLA Results



facebook Home Profile Friends Inbox 1 Ying Zhao Settings Logout Search

Advertising Pages Share Connect

## Facebook Advertising

Reach your exact audience and connect real customers to your business.

or manage your existing ads

Overview Prepare Step By Step Case Studies

### 2. Targeting

Location: United States

☒ Everywhere

☐ By State/Province

☐ By City

Age: 18 - Any

#### Connect with Real People

- Reach over 300,000,000 active Facebook users.
- Attach social actions to your ads to increase relevance.
- Create demand for your product with relevant ads.

#### Create Your Facebook Ad

- Quickly create image and text-based ads.
- Advertise your own web page or something on Facebook like a Page or an Event.
- Choose to pay per click (CPC) or impression (CPM) in your local currency.

#### Optimize Your Ads

- Track your progress with real-time reporting.
- Gain insight about who's clicking on your ad.
- Make modifications to maximize your results.

Free ice cream

Hope you enjoy free ice cream on Election Day! Be a Ben & Jerry's fan for the latest updates and the first to know about

Female	Male	Age
		13-17
		18-24
		25-34
		35-44
		45+

Ads can be directed to those who match required characteristics



# Create a dedicated application



LLA for Navy Recruiting | Facebook - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Forward Stop Reload Home Search Favorites

Address http://www.facebook.com/apps/application.php?id=127300506246#/apps/application.php?v=wall&id=127300506246

Search the web Search Images Weather News Maps Resize Highlight

Web Search Bookmarks Settings Mail Weather Flickr Finance eBay

LLA for Navy Recruiting | Faceb... Add Tab

facebook

Keep me logged in Forgot your password? Login

Sign Up LLA for Navy Recruiting is on Facebook

Sign up for Facebook to connect with LLA for Navy Recruiting.

LLA for Navy Recruiting

Wall Info Reviews Discussions

LLA for Navy Recruiting Just Fans

Go to Application

Information

Category All

This application was not developed by Facebook.

Contact Developer Report Application

LLA for Navy Recruiting has no more posts.

- Post marketing information disseminated to fans
- Grow the fan space



# Ads Placed via Facebook to Matched Profiles & Response Results Tracked





Suggest to Friends  
Subscribe via SMS

Honor, Courage and Commitment.  
These are the Core Values of the Navy, and as a Sailor they will be your core values too.


**Information**

Location:  
103-45 lefferts blvd  
Richmond Hill, NY, 11419

Phone:  
718-843-4380

**Fans**

6 of 36 fans See All



NAVY RECRUITING STATION RICHMOND HILL NY

Wall Info Photos Discussions Reviews Boxes

Write something...

Attach:  Share

NAVY RECRUITING STATION RICHMOND HILL NY Just Fans

NAVY RECRUITING STATION RICHMOND HILL NY alright we are set up!  
July 21 at 5:45am · Comment · Like

NAVY RECRUITING STATION RICHMOND HILL NY you know the drill

**DEP Meeting**  
Time: 4:00PM Thursday, August 6th  
Location: 103-45 lefferts blvd  
July 21 at 5:32am · Comment · Like · Share

Frank Snaith likes this.

NAVY RECRUITING STATION RICHMOND HILL NY paintball?  
July 21 at 5:33am · Report

Write a comment...

**RECENT ACTIVITY**

NAVY RECRUITING STATION RICHMOND HILL NY discussed getting in shape on the NAVY RECRUITING STATION RICHMOND HILL NY discussion board.

NAVY RECRUITING STATION RICHMOND HILL NY edited their Phone, Location and Website.

Create an Ad

**Custom Wall Murals** ×

  
MuralsYourWay.com has over 5,000 images or send us yours and we'll create a wall mural in any size! As seen on HGTV & Extreme Makeover.  
Like

**T-Mobile BlackBerry** ×

  
Find out first when the new BlackBerry® Bold™ 9700 with Wi-Fi calling arrives from T-Mobile™. Become a fan to find out more.  
Become a Fan

**Donate a Car for BREAST CANCER** ×

  
**FAST EASY DONATION**  
Free Towing & Tax Deduction.  
No Costs :- Running or Not  
Top Breast Cancer Charity  
Please Help Call  
888-773 3784



# Facebook Profile Data Example



**▼ Connect**

**Interests (36):**  
 adventure, archipelagoes, cavalry, cold war, dragons, fantasy, fiction, geography, heraldry, history, immortality, longevity, love story, magic, military history, mindpowers, mounted combat, nanotechnology, nanowrimo, national novel writing month, naval history, novel writing, parapsy, psionics, psychic powers, regeneration, sailing ships, science fiction, storytelling, time travel, tropical islands, warfare, wizardry

**External Services:**

tanelos@livejournal.com  
 rizanabsith  
 191317559  
 rizanabsith  
 rizanabsith@hotmail.com  
 rizanabsith@gotrillian.com  
 banazir  
 banazir

**Schools:**  
 Hillsborough Elementary School - Hillsborough, NJ (1983 - 1984)  
 Scott Lake Elementary School - Lakeland, FL (1984 - 1985)  
 Severna Park Middle School - Severna Park, MD (1985 - 1987)  
 Severn School - Severna Park, MD (1987 - 1989)  
 Johns Hopkins University - Baltimore, MD (1989 - 1993)  
 University of Illinois at Urbana-Champaign - Urbana, IL (1993 - 1998)

**Friends [View Entries]**

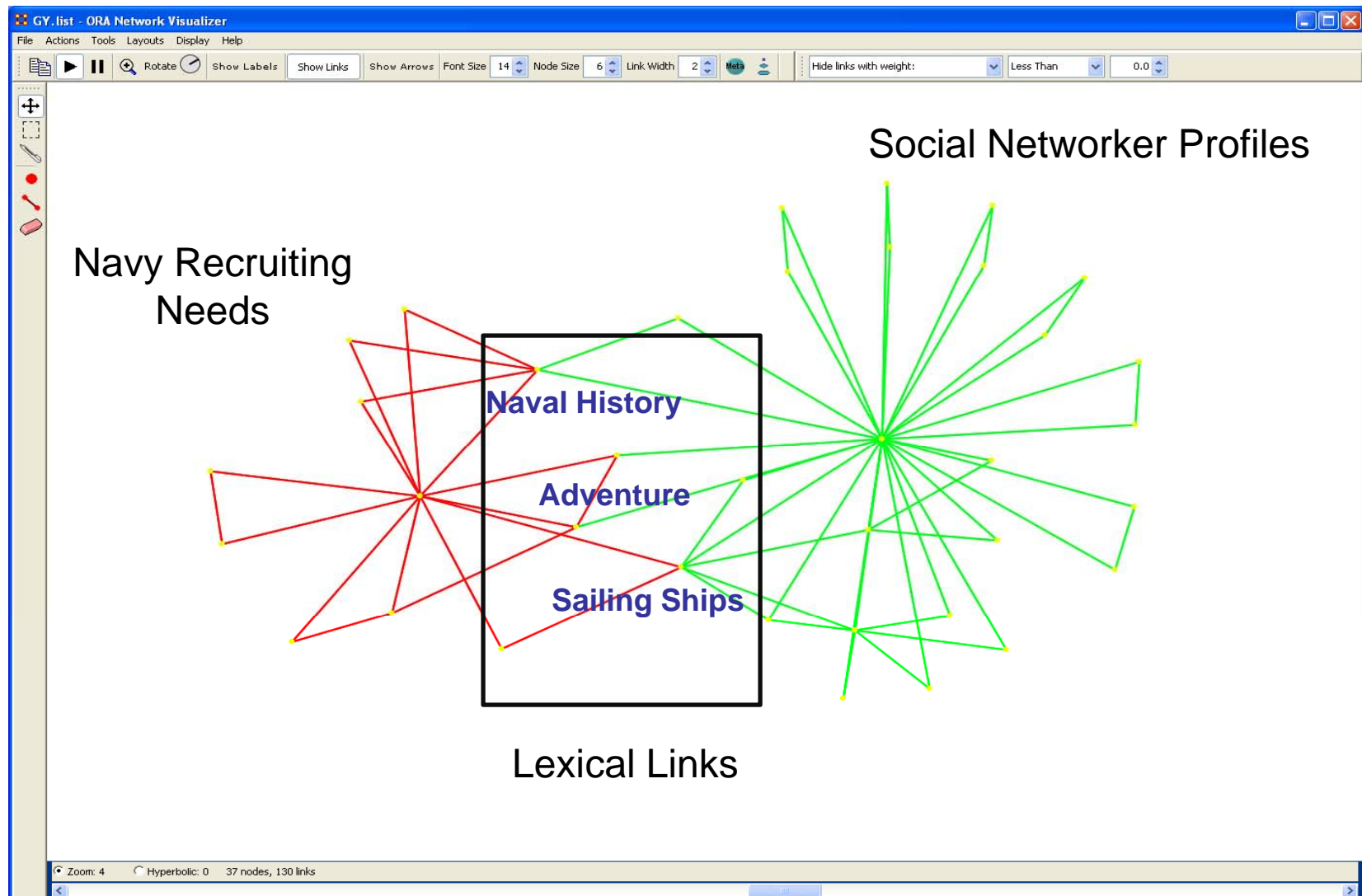
**▼ Friends (66):**  
 amishlurker, andrewwyldgonzo, anim8t0rboi, banazir, bdm7935, burkhardt, carida\_46, cavlec, cheyinka, cloverzj, cretaceousrick, darana, deire, dragnflye, dsthenes, elvenwanderer, figgylicious, gngr, gondhir, gothic\_spire, illusio, istari\_ala, jadzidax, kaladhwen, klawz, koinonia, ldymlissa, masaga, masteralida, megruder, mijesbar, mirabehn, miyeko, mom\_counsel, mrowe, myng\_rabbyt, narvi, neoookami, nobuddy69, pathofthejus, queenmidalah, ringlass, rivendellrose, sahtyinepu, smegalicious, sperose, sui\_degeneris, tabula\_xrasa, taiji\_jian, tamf, teremala, the\_wanlorn, t, twinothugin, yodge, zengeneral, zurich31

2. Data		
<input type="checkbox"/> w1dataset.xls	Restricted Access	wave 1 general attributes
MS Excel - 314 KB - 24 downloads		
<input type="checkbox"/> w1fbk.net	Restricted Access	wave 1 facebook friends
Plain Text - 1 MB - 21 downloads		
<input type="checkbox"/> w1hsgrp.net	Restricted Access	wave 1 housing groupmates
Plain Text - 218 KB - 21 downloads		
<input type="checkbox"/> w1pic.net	Restricted Access	wave 1 picture friends
Plain Text - 357 KB - 21 downloads		
<input type="checkbox"/> w1room.net	Restricted Access	wave 1 roommates
Plain Text - 170 KB - 21 downloads		
<input type="checkbox"/> w1tastes.xls	Restricted Access	wave 1 cultural tastes
MS Excel - 1 MB - 21 downloads		



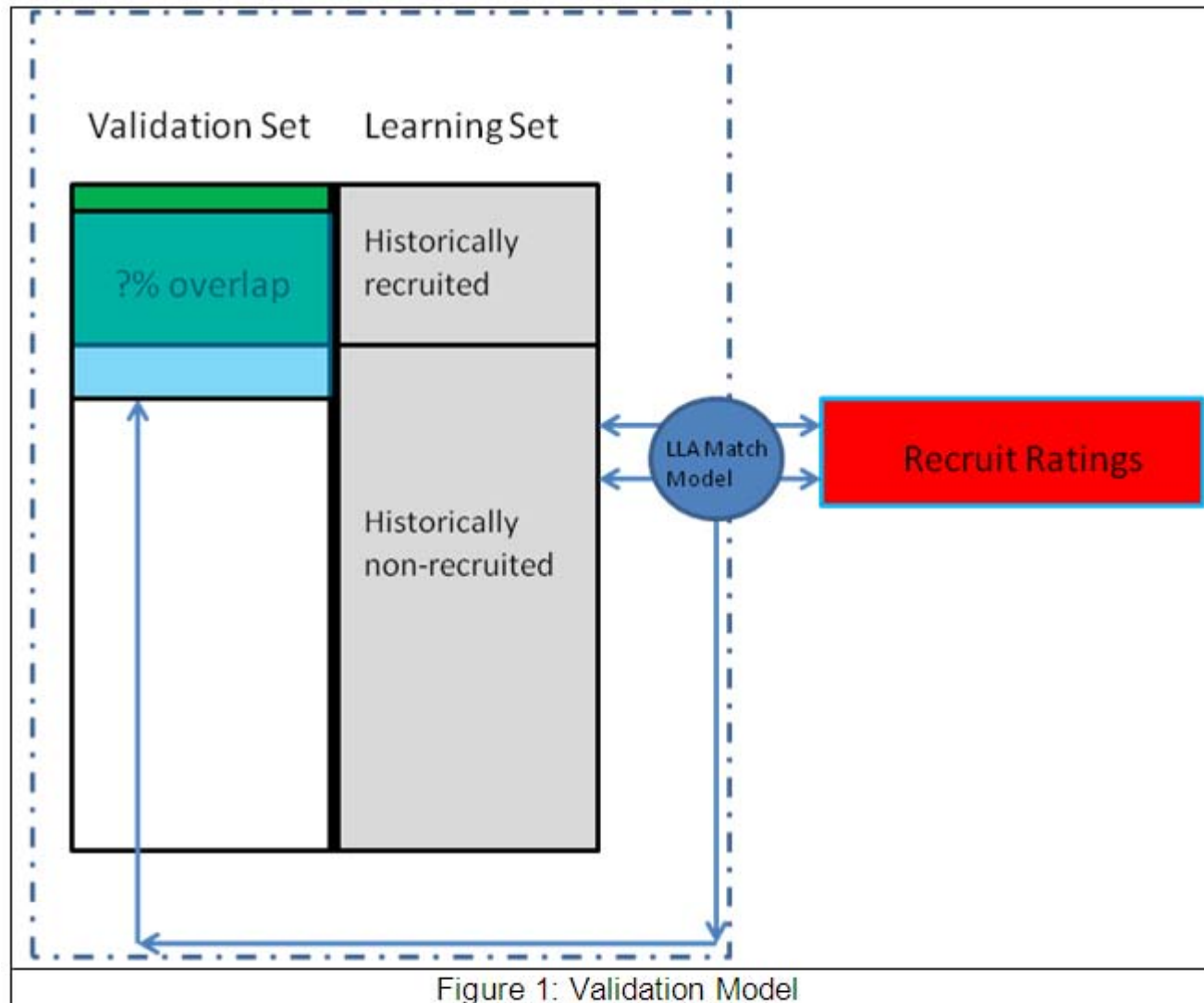


# Notional Lexical Link Map Matching Profiles to Navy Ratings





# LLA Recruiting Validation Methods







# LLA Recruiting Validation Method



- **Step 1:** Consider historical sample profiles of recruited personnel, mixed with general populations, for example, samples of youth population in US. The profiles should be similar to the self-reported education and personality characteristics from Facebook.
- **Step 2:** Divide the samples into a validation and a learning set. Set aside the validation set. Use the learning set to match the recruit ratings with the proposed LLA and generate a LLA Match Model.
- **Step 3:** Apply the LLA Match Model to the validation set to recommend the top candidates. Compute the percentage of overlap between the recommended recruits with those who were successfully selected historically (by experts). (Please see the process illustration in Figure 1.) If the overlap percentage or correlation is high, we will consider the matching method as valid.
  - This method can then be used dynamically in automation and thus reduce required manpower.



# Rating Descriptions vs. Population Segment



- Data Used
  - Enlisted ratings (AB.doc, AC.doc, AD.doc....)
    - Descriptions
  - JAMRS\_2008\_SegmentationGuide.pdf
    - 66 segments for 16-20 years of age
- For Validation
  - FY09 National Segmentation Profile.xls
    - Accession rates for all services, for all segments were used for correlation



# Result: Lexical Link Table



- **Rows:** Segment description ids
- **Columns:** Enlisted rating ids
- **Number in the cells:** the number of matched word/term categories for a segment and a rating.

Enlisted ratings

Segments

	A	B	C	D	E	F	G	H
		Segment01.txt	Segment02.txt	Segment03.txt	Segment04.txt	Segment05.txt	Segment06.txt	Segment07.txt
1								
2	AB.doc					1	1	
3	AC.doc					1	1	
4	AD.doc					1	1	
5	AECF.doc		1	1	1	2	2	2
6	AG.doc		1			1	2	1
7	AIRCREW.doc							
8	AM.doc					1	1	
9	AME.doc					1	1	
10	AQ.doc					1	1	
11	AS.doc					1	1	
12	AV.doc					1	1	
13	AZ.doc					1	1	
14	Admin Job Family.doc					1	1	
15	Aviation Mech Job Family.doc					1	1	
16	Avt Deck Job Family.doc					1	1	
17	BM.doc					1	1	
18	BU.doc					1	2	2
19	CE.doc					1	1	
20	CM.doc					1	2	2
21	CS.doc					1	1	
22	CSS.doc					1	1	
23	CTI HARDCARD JAN08.doc							
24	CTM.doc							
25	CTN HARDCARD JAN08.doc							
26	CTR.doc					1	1	
27	CTT AEF HARDCARD JAN08.doc					1	1	
28	CTT HARDCARD JAN08.doc					1	1	
29	DC.doc		1	1	1	2		1
30	EA.doc					1	1	
31	EL4 Job Family.doc					1	1	
32	EM.doc		1	1	1	2		1
33	EN.doc					1	1	
34	EO.doc		1	1	1	2		1
35	EOD.doc							
36	GM.doc		1	1	1	2		1
37	GSE.doc		1	1	1	2		1
38	GSM.doc		1	1	1	2		1
39	HM.doc		1	1	2	1		
40	HT.doc					1	1	
41	IC.doc		1	1	1	2		1
42	IS.doc							
43	IT.doc					1	1	
44	LS.doc					1	1	
45	LSSS.doc					1	1	
46	MA.doc					1	1	
47	MC.doc					1	1	
48	MM SS.doc					1	1	
49	MM.doc			1	1	2		1
50	MN.doc					1	1	
51	MR.doc					1	1	
52	MT.doc							
53	ML.doc							

**Number in the cells:** the number of matched word/term categories for a segment and a rating.

Enlisted ratings

Segments

	AF	AG	AH	AI	AJ	AK	
1	Segment31.txt	Segment32.txt	Segment33.txt	Segment34.txt	Segment35.txt	Segment36.txt	Segment
42	IS.doc			1			
43	IT.doc		1			1	1
44	LS.doc		1	1		1	1
45	LSSS.doc		1	1		1	1
46	MA.doc		1	1		1	1
47	MC.doc		1	1		1	1
48	MMSS.doc		1	1		1	1
49	MM.doc		2	2	1	1	1
50	MN.doc		1	1		1	1
51	MR.doc		1	1		1	1
52	MT.doc			1			
53	MU.doc						
54	ND.doc				1	1	
55	Nuke Power.doc						
56	OS.doc		1			1	1
57	PC.doc			1			
58	PR.doc		1	1		1	1
59	PS.doc		1	1		1	1
60	QM.doc		2	1	1	1	1
61	RP.doc		1	1		1	1
62	SB.doc		1	1	1		
63	SECF.doc		2	2	1	1	2
64	SENG Job Family.doc		1			1	1
65	SH.doc					1	1
66	SK.doc					1	1
67	SKSS.doc					1	1
68	SN.doc						
69	SO.doc				2		1
70	STG AEF.doc				1	1	1
71	STG.doc					1	1
72	SW.doc					1	1
73	Seabee UCT.doc						
74	UT.doc		1			1	1
75	YN.doc		1	1		1	1
76	YNSS.doc		1	1		1	1
77	# of Word Categories Linked	2	7	8	4	4	4
78	LANGUAGE--SPANISH,SNIPING	PROGRAMMING--TELEVISION,OPERATING	VOCATIONAL--TECHNICAL,CERTIFICATE	LANGUAGE--SPANISH,SNIPING	PROGRAMMING--TELEVISION,OPERATING	LANGUAGE--SPANISH,SNIPING	PROGR/TELEVIS
79	PROGRAMMING--TELEVISION,OPERATING	DEVELOPMENT--SKILL,CAREER	PROGRAMMING--TELEVISION,OPERATING	2ND--CITY,LOCATED,CREDITS	SERVE--PEOPLE,BATTLEFIELD	DEVELOPMENT--SKILL,CAREER	SERVE--
80		UPPER--WHITE,HOURS	SERVE--PEOPLE,BATTLEFIELD	INTERESTED--FINANCIAL,PURSURING	2ND--CITY,LOCATED,CREDITS	INTERESTED--FINANCIAL,PURSURING	UPPER--
81		INTERESTED--MECHANICS,PURSURING	UPPER--WHITE,HOURS	SELF--DISCIPLINE,PACED	DEVELOPMENT--SKILL,CAREER	TECHNOLOGY--EDGE,AAS	INTERES
82		SELF--DISCIPLINE,PACED	INTERESTED--SUPPORTING,PURSURING				COLLEG
83		TECHNOLOGY--EDGE,AAS	SELF--DISCIPLINE,PACED				
84		COLLEGE--MONEY,THOMAS	TECHNOLOGY--EDGE,AAS				
85	Word Categories Linked		COLLEGE--MONEY,THOMAS				

# of word categories linked: unique # of categories, i.e. 2, across all the ratings for Segment31

2	ALL SVCS Accessions Rate	0.003824919	0.00706807	0.006161745	0.005987297	0.008384598	0.009651198	0.00588802
3	# of Word Categories Linked	2	2	2	3	6	5	2
4		PROGRAMMING-- READ, OPERATING	SERVE-- PEOPLE, BATTLEFIELD	SERVE-- PEOPLE, BATTLEFIELD	SERVE-- PEOPLE, BATTLEFIELD	PART--WORK		
5		INTERESTED-- MECHANICS, PURSUING	SELF-- DISCIPLINE, PACED	SELF-- DISCIPLINE, PACED	DEVELOPMENT-- SKILL, CAREER	PROGRAMMING-- READ, OPERATING		
6					INTERESTED-- MECHANICS, PURSUING	DEVELOPMENT-- SKILL, CAREER		
7						INTERESTED-- MECHANICS, F		
8						SELF--DISCIPLINE		
9						TECHNOLOGY		
10								
11	Word Categories Linked							
12		Correlation						
13		0.435522936						
14								
15								
16								
17								
18								
19								
20								
21								
22								
23								
24								
25								
26								
27								

- Pearson correlation between “# of Word Categories Linked” and the accession rate for “ALL SVCS Accessions” from FY09 National Segmentation Profile.xls.
- The correlation of Row 2 vs. 3 is 0.43 (on a scale of 0 to 1)
- Provides Initial validation for this methodology

Excel Help

pearson Search

Syntax

PEARSON(array1,array2)

Array1 is a set of independent values.

Array2 is a set of dependent values.

Remarks

- The arguments must be either numbers or names, array constants, or references that contain numbers.
- If an array or reference argument contains text, logical values, or empty cells, those values are ignored; however, cells with the value zero are included.
- If array1 and array2 are empty or have a different number of data points, PEARSON returns the #N/A error value.
- The formula for the Pearson product moment correlation coefficient,  $r$ , is:

$$r = \frac{\sum (x - \bar{x})(y - \bar{y})}{\sqrt{\sum (x - \bar{x})^2 \sum (y - \bar{y})^2}}$$



# Conclusions

- Serves as initial validation LLA method
- We expect that more detailed, individual profiles (e.g. facebook) will increase this correlation
- Improve the accession rate in real-life
- Word/term matches can be used to fit different segment youth to different ratings



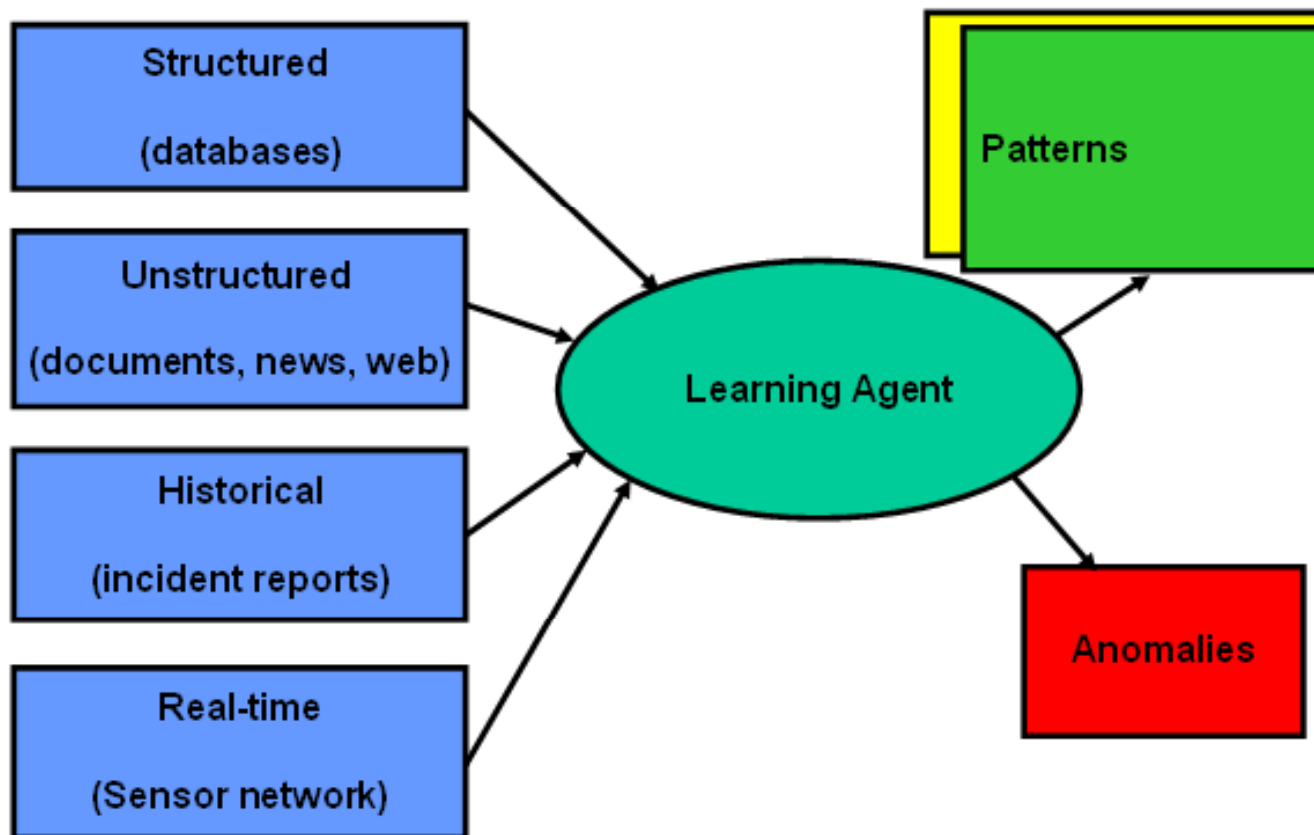
# References





# Collaborative Learning Agents (CLA)

(<http://www.quantumii.com>)





# Internet References



- <http://blog.jobzcafe.com/2008/09/08/facebook-for-recruiting-whats-working/>
- [http://www.facebook.com/pages/US-Navy-SEAL-SWCC-EOD-Diver-Rescue-Swimmer-recruiting/166379650511?v=feed&story\\_fbid=247595570511&ref=mf](http://www.facebook.com/pages/US-Navy-SEAL-SWCC-EOD-Diver-Rescue-Swimmer-recruiting/166379650511?v=feed&story_fbid=247595570511&ref=mf)
- <http://www.facebook.com/pages/Chambersburg-PA/Navy-Recruiting-Station-Chambersburg-PA/116537271249>